

PDF reweighting in the Hessian matrix approach

Hannu Paukkunen^{a,b} and Pia Zurita^c

^a*Department of Physics, University of Jyväskylä, P.O. Box 35, FI-40014, Finland*

^b*Helsinki Institute of Physics, University of Helsinki, P.O. Box 64, FI-00014, Finland*

^c*Departamento de Física de Partículas and IGFAE, Universidade de Santiago de Compostela, E-15782 Galicia, Spain*

E-mail: hannu.paukkunen@jyu.fi, pia.zurita@usc.es

ABSTRACT: We introduce the Hessian reweighting of parton distribution functions (PDFs). Similarly to the better-known Bayesian methods, its purpose is to address the compatibility of new data and the quantitative modifications they induce within an existing set of PDFs. By construction, the method discussed here applies to the PDF fits that carried out a Hessian error analysis using a certain $\Delta\chi^2$ tolerance. The principle is validated by considering a simple, transparent example. We are also able to establish an agreement with the Bayesian technique provided that the $\Delta\chi^2$ criterion is appropriately accounted for and that a purely exponential Bayesian likelihood is assumed. As a practical example, we discuss the inclusive jet production at the LHC.

Contents

1	Introduction	1
2	The Hessian method	2
3	The Hessian reweighting	4
4	Bayesian methods	6
5	Simple example	9
6	Equivalence of the Bayesian and Hessian reweighting	12
7	Inclusive jet production at the LHC	13
8	Summary	18

1 Introduction

A large part of the present high-energy collider physics depends, in one way or another, on the knowledge of the parton distribution functions (PDFs). The use of PDFs leans on a cornerstone theorem of Quantum Chromo Dynamics (QCD), the collinear factorization [1, 2]. Although it can be formally proven only for the most simple cases, it is often assumed to work in general. Ultimately, it is the agreement with the experimental data that decides whether such assumption is valid.

The PDFs are traditionally determined in global analyses [3–6] finding the parameters that can optimally reproduce a variety of experimental data. This is a complex procedure requiring the ability to efficiently solve the parton evolution equations and to calculate higher-order QCD cross-sections. Nontrivial issues are also the actual way of finding the best fit and quantifying its uncertainties. Although various PDF parametrizations are publicly available for a general user, for a long period of time it was difficult for e.g. an experimental collaboration to understand what would be the implications of their measurements in the context of a global PDF fit. For example, although a given measurement would be known to be most sensitive to, say, the up quark distribution, some other data in a global fit may already provide the up quarks with more stringent constraints, and the real advantage of the measurement would be due to a sub-leading contribution, say, from the strange quarks.

The Bayesian reweighting technique, first introduced in [7] and later on elaborated by the Neural Network PDF (NNPDF) collaboration [8, 9], provides a way of addressing the consistency and quantitative effects of a new experimental evidence in terms of PDF fits.

In essence, the underlying probability distribution, represented in the NNPDF philosophy [6] by an ensemble (~ 1000) of PDF replicas, is updated by assigning each replica a certain weight based on how well it reproduces the new data. This method has become an increasingly popular way to estimate the effects of e.g. new LHC measurements [10–16]. However, the majority of the existing PDF fits use a rather different way of quantifying the PDFs and their uncertainties. Along with the best fit found by χ^2 minimization, it is customary to provide a collection (~ 50) of Hessian error sets [17] that quantify the neighborhood of the central fit within a certain confidence criterion $\Delta\chi^2$. An extension of the Bayesian reweighting technique to this particular case was suggested in [18], and has thereafter been used in some occasions [11, 19, 20]. However, a recent study [21] revealed some imperfections when comparing the results from reweighting to the ones obtained by a direct fit done in the usual way.

Here, we take a different strategy. Based on the ideas presented in Ref. [22], our principal goal is to show how a general user can directly study the consistency and consequences of a new data set within an existing set of PDFs that comes with the Hessian error sets without having to generate multitude of PDF replicas. The method naturally incorporates the confidence criterion $\Delta\chi^2$ defined in the original fit. By considering a simple numerical example, we argue that the method is perfectly compatible with a new fit. Our second objective is to understand how the procedure suggested here relates to the Bayesian reweighting and to the discrepancies found in [21]. We prove that the original Bayesian method, proposed in [7] and advocated recently in [23], is equivalent with the one introduced here once the $\Delta\chi^2$ criterion is properly incorporated.

2 The Hessian method

The usual definition of an optimum correspondence between data and a set of PDFs $f \equiv f(x, Q^2)$ that depends on certain fit parameters $\{a\}$, is the minimum of a χ^2 -function. In its most simple form, we can write it as

$$\chi^2\{a\} = \sum_k \left[\frac{X_k^{\text{theory}}[f] - X_k^{\text{data}}}{\delta_k^{\text{data}}} \right]^2, \quad (2.1)$$

where $X_k^{\text{theory}}[f]$ are the theory predictions depending on the PDFs. The corresponding experimental values are denoted by X_k^{data} and their uncertainty by δ_k^{data} . Modifications to this definition are necessary if the experimental errors are correlated or some data sets are emphasized in the fit by assigning them an additional weight. In the Hessian approach to quantify the PDF errors [17], the behaviour of χ^2 around the best fit S_0 is approximated by a second order polynomial in the space of fit parameters $\{a\}$

$$\chi^2\{a\} \approx \chi_0^2 + \sum_{ij} \delta a_i H_{ij} \delta a_j, \quad (2.2)$$

where $\delta a_j \equiv a_j - a_j^0$ are the excursions from the best-fit values and χ_0^2 is the minimum value of χ^2 . Being symmetric, the Hessian matrix H_{ij} has N_{eig} orthonormal eigenvectors

$\mathbf{v}^{(k)}$ and eigenvalues ϵ_k satisfying

$$H_{ij}v_j^{(k)} = \epsilon_k v_i^{(k)}, \quad (2.3)$$

$$\sum_j v_j^{(k)} v_j^{(\ell)} = \sum_j v_k^{(j)} v_\ell^{(j)} = \delta_{k\ell}. \quad (2.4)$$

Defining a new set of variables as

$$z_k \equiv \sqrt{\epsilon_k} \sum_j v_j^{(k)} \delta a_j, \quad (2.5)$$

one easily finds that

$$\chi^2\{a\} \approx \chi_0^2 + \sum_i z_i^2. \quad (2.6)$$

That is, the transformation in Eq. (2.5) diagonalizes the Hessian matrix. At this point, a $\Delta\chi^2$ criterion is needed to specify how much the term $\sum_i z_i^2$ can grow while the corresponding PDFs still remain “acceptable”. Those PDF fits that employ the ideal choice $\Delta\chi^2 = 1$ are usually limited to a subset of world data [24, 25], while the global fits prefer to take $\Delta\chi^2 > 1$ [4, 5] to allow inconsistencies among different data sets. It follows that the corresponding uncertainty for a PDF-dependent quantity $\mathcal{O} = \mathcal{O}[f]$ can be computed as

$$(\Delta\mathcal{O})^2 = \Delta\chi^2 \sum_k \left(\frac{\partial\mathcal{O}}{\partial z_k} \right)^2. \quad (2.7)$$

An essential feature of the Hessian approach is the introduction of the PDF error sets S_k^\pm , defined customarily (along with the best fit S_0) in the z -space as

$$\begin{aligned} z(S_0) &= (0, 0, \dots, 0), \\ z(S_1^\pm) &= \pm \sqrt{\Delta\chi^2} (1, 0, \dots, 0), \\ z(S_2^\pm) &= \pm \sqrt{\Delta\chi^2} (0, 1, \dots, 0), \\ &\vdots \\ z(S_{N_{\text{eig}}}^\pm) &= \pm \sqrt{\Delta\chi^2} (0, 0, \dots, 1). \end{aligned} \quad (2.8)$$

Using these sets, one can evaluate the derivatives in Eq. (2.7) by a linear approximation

$$\left(\frac{\partial\mathcal{O}}{\partial z_k} \right) \approx \frac{\mathcal{O}[S_k^+] - \mathcal{O}[S_k^-]}{2\sqrt{\Delta\chi^2}}, \quad (2.9)$$

such that

$$(\Delta\mathcal{O})^2 = \frac{1}{4} \sum_k (\mathcal{O}[S_k^+] - \mathcal{O}[S_k^-])^2. \quad (2.10)$$

This formula, or a generalization for asymmetric errors, provides an extremely simple and useful recipe for propagating the PDF-uncertainties to observables.

3 The Hessian reweighting

Let us now consider a new set of data $\vec{y} = y_1, y_2, \dots, y_{N_{\text{data}}}$ with covariance matrix C . Our goal here is twofold: to find out whether these new data are consistent within the original set of PDFs and, if so, what would be the effect of incorporating them into the original analysis. In order to answer these questions, we consider a function χ_{new}^2 defined as

$$\chi_{\text{new}}^2 \equiv \chi_0^2 + \sum_k^{N_{\text{eig}}} z_k^2 + \sum_{i,j=1}^{N_{\text{data}}} (y_i[f] - y_i) C_{ij}^{-1} (y_j[f] - y_j), \quad (3.1)$$

where we have simply added the contribution of the new data on top of the “old” χ^2 in Eq. (2.6). Using a similar linear approximation as earlier, we can estimate the theoretical values $y_i[f]$ in arbitrary z -space coordinates by

$$y_i[f] \approx y_i[S_0] + \sum_{k=1}^{N_{\text{eig}}} \left. \frac{\partial y_i[S]}{\partial z_k} \right|_{S=S_0} z_k \approx y_i[S_0] + \sum_{k=1}^{N_{\text{eig}}} D_{ik} w_k, \quad (3.2)$$

where we have defined

$$D_{ik} \equiv \frac{y_i[S_k^+] - y_i[S_k^-]}{2} \quad (3.3)$$

$$w_k \equiv \frac{z_k}{\sqrt{\Delta\chi^2}}. \quad (3.4)$$

Thus, χ_{new}^2 is a continuous, quadratic function of the parameters w_k , and its minimum is given by

$$\vec{w}^{\text{min}} = -\mathbf{B}^{-1} \vec{a}, \quad (3.5)$$

where the matrix \mathbf{B} and vector \vec{a} are

$$B_{kn} = \sum_{i,j} D_{ik} C_{ij}^{-1} D_{jn} + \Delta\chi^2 \delta_{kn}, \quad (3.6)$$

$$a_k = \sum_{i,j} D_{ik} C_{ij}^{-1} (y_j[S_0] - y_j). \quad (3.7)$$

An important feature of the solution is the “penalty term”

$$P \equiv \Delta\chi^2 \sum_{k=1}^{N_{\text{eig}}} (w_k^{\text{min}})^2, \quad (3.8)$$

which we can use to decide whether the new data set is consistent within the original PDFs. First, if $P \ll \Delta\chi^2$ the new data could be incorporated into the original fit without causing a conflict with the other data. On the other hand, if $P \gtrsim \Delta\chi^2$ the new data appears to display tension with the considered set of PDFs. This, however, does not necessarily mean that the new data would be incompatible with the other data. A situation like this may arise if the new data probe unconstrained components of PDFs whose behaviour was fixed by hand. For example, some recent PDF fits [26] still assume $s(x) \propto (u(x) + d(x))$ for the

strange quark distribution and confronted with data sensitive to the strange quarks could lead to this kind of situation.

The components of the weight vector \vec{w}^{\min} also specify the set of PDFs f^{new} that corresponds to the new global minimum. They can be easily calculated by taking $y_i = f(x, Q^2)$ in Eq. (3.2). That is,

$$f^{\text{new}} \approx f_{S_0} + \sum_{k=1}^{N_{\text{eig}}} \left(\frac{f_{S_k^+} - f_{S_k^-}}{2} \right) w_k^{\min}. \quad (3.9)$$

The resulting new PDFs are linear combinations of the original ones — they have been “reweighted”. We note that the new PDFs constructed in this way still satisfy the necessary sum rules. For instance, as the original best fit S_0 and the error sets S_k^{\pm} satisfy the momentum sum rule

$$\int_0^1 dx \sum_f f_{S_0} = \int_0^1 dx \sum_f f_{S_k^{\pm}} = 1, \quad (3.10)$$

then

$$\begin{aligned} \int_0^1 dx \sum_f f^{\text{new}} &= \int_0^1 dx \sum_f f_{S_0} + \sum_k \frac{w_k^{\min}}{2} \left[\int_0^1 dx \sum_f f_{S_k^+} - \int_0^1 dx \sum_f f_{S_k^-} \right] \\ &= 1 + \sum_k \frac{w_k^{\min}}{2} [1 - 1] = 1. \end{aligned}$$

Due to the linearity of the parton evolution equations, also f^{new} satisfies them. Thus, the reweighted distributions comprise a proper set of PDFs which can be consistently utilized in pQCD calculations. Furthermore, assuming that the original confidence criterion is not much altered by the inclusion of the new data, one can also construct the new PDF error sets. Indeed, Eq. (3.1) can be rewritten as

$$\chi_{\text{new}}^2 = \chi_{\text{new}}^2 \big|_{\vec{w}=\vec{w}^{\min}} + \sum_{ij} \delta w_i B_{ij} \delta w_j, \quad (3.11)$$

where $\delta \vec{w} = \vec{w} - \vec{w}^{\min}$, and the matrix \mathbf{B} takes the role of the Hessian matrix (compare to Eq. (2.2)). This can be brought into a diagonal form by an analogue of the transformation in Eq. (2.5) and the new error sets defined exactly as was done in Eq. (2.8).

Non-linear extension of the Hessian reweighting

The linear approximation of Eq. (3.2) can be improved by including also quadratic terms of w_k as

$$\begin{aligned} y_i[S] &= y_i[S_0] + \sum_{k=1}^{N_{\text{eig}}} \left(\frac{y_i[S_k^+] - y_i[S_k^-]}{2} \right) w_k \\ &\quad + \sum_{k=1}^{N_{\text{eig}}} \left(\frac{y_i[S_k^+]}{2} + \frac{y_i[S_k^-]}{2} - y_i[S_0] \right) w_k^2, \end{aligned} \quad (3.12)$$

correcting for the possible non-linear behaviour $y_i [S_k^+] - y_i [S_0] \neq y_i [S_0] - y_i [S_k^-]$. Using this in Eq. (3.1), χ_{new}^2 becomes a quartic function of w_k and its minimum can be found using numerical methods. The corresponding PDFs can be computed by taking $y_i = f(x, Q^2)$ in Eq. (3.12). The matrix \mathbf{B} in Eq. (3.11) gets replaced by

$$B_{kn}^{\text{non.lin}} = \sum_{i,j} \left(\frac{\partial y_i[f]}{\partial w_k} \right) C_{ij}^{-1} \left(\frac{\partial y_j[f]}{\partial w_n} \right) + \sum_{i,j} \left(\frac{\partial^2 y_i[f]}{\partial w_k \partial w_n} \right) C_{ij}^{-1} (y_j[f] - y_j) + \Delta \chi^2 \delta_{kn}, \quad (3.13)$$

where the partial derivatives read

$$\left(\frac{\partial y_i[f]}{\partial w_k} \right) = \left(\frac{y_i [S_k^+] - y_i [S_k^-]}{2} \right) + 2 \left(\frac{y_i [S_k^+]}{2} + \frac{y_i [S_k^-]}{2} - y_i [S_0] \right) w_k \quad (3.14)$$

$$\left(\frac{\partial^2 y_i[f]}{\partial w_k \partial w_n} \right) = 2 \left(\frac{y_i [S_k^+]}{2} + \frac{y_i [S_k^-]}{2} - y_i [S_0] \right) \delta_{kn}, \quad (3.15)$$

and are understood to be evaluated at the found minimum.

4 Bayesian methods

Given a large ensemble of PDFs f_k , $k = 1 \dots N_{\text{rep}}$, such as those of the NNPDF collaboration [6], that represents the underlying probability distribution $\mathcal{P}_{\text{old}}(f)$ of the PDFs, one can compute the expectation value $\langle \mathcal{O} \rangle$ and variance $\delta \langle \mathcal{O} \rangle$ for an observable \mathcal{O} as

$$\langle \mathcal{O} \rangle = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \mathcal{O} [f_k], \quad (4.1)$$

$$\delta \langle \mathcal{O} \rangle = \sqrt{\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} (\mathcal{O} [f_k] - \langle \mathcal{O} \rangle)^2}. \quad (4.2)$$

Using the laws of statistics, the initial probability distribution $\mathcal{P}_{\text{old}}(f)$ can be updated to include also additional information contained in a new set of data \vec{y} , since, by the Bayes theorem,

$$\mathcal{P}_{\text{new}}(f) \propto \mathcal{P}(\vec{y}|f) \mathcal{P}_{\text{old}}(f), \quad (4.3)$$

where $\mathcal{P}(\vec{y}|f)$ stands for the conditional probability (the likelihood function) for the new data, given a set of PDFs. It follows that the average value for any observable depending on the PDFs becomes a weighted average

$$\langle \mathcal{O} \rangle_{\text{new}} = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \omega_k \mathcal{O} [f_k], \quad (4.4)$$

$$\delta \langle \mathcal{O} \rangle_{\text{new}} = \sqrt{\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \omega_k (\mathcal{O} [f_k] - \langle \mathcal{O} \rangle_{\text{new}})^2}, \quad (4.5)$$

where the weights ω_k turn out to be proportional to the likelihood function $\mathcal{P}(\vec{y}|f)$. The question of how to choose the likelihood appropriately has been recently addressed in [23] but the conclusive answer, if it exists, remains to be given. Two options, corresponding to different choices of the likelihood, have been discussed in the literature. The one suggested originally by Giele and Keller (GK) [7] follows from taking $\mathcal{P}(\vec{y}|f)d^n y$ as the probability to find the new data to be confined in a differential element $d^n y$ around \vec{y} resulting in

$$\omega_k^{\text{GK}} = \frac{\exp[-\chi_k^2/2]}{(1/N_{\text{rep}}) \sum_{k=1}^{N_{\text{rep}}} \exp[-\chi_k^2/2]}, \quad (4.6)$$

where

$$\chi_k^2 = \sum_{i,j=1}^{N_{\text{data}}} (y_i[f_k] - y_i) C_{ij}^{-1} (y_j[f_k] - y_j). \quad (4.7)$$

The option advocated by the NNPDF collaboration derives from taking $\mathcal{P}(\vec{y}|f)d\chi$ as the probability for the corresponding χ^2 to be confined in a differential volume $d\chi$ around χ , giving instead¹

$$\omega_k^{\text{chi-squared}} = \frac{(\chi_k^2)^{(N_{\text{data}}-1)/2} \exp[-\chi_k^2/2]}{(1/N_{\text{rep}}) \sum_{k=1}^{N_{\text{rep}}} (\chi_k^2)^{(N_{\text{data}}-1)/2} \exp[-\chi_k^2/2]}, \quad (4.8)$$

which was shown to be consistent with a direct fit in the NNPDF framework [8, 9]. It was pointed out in Ref. [23] that the former weights contain more information on the new data than the latter ones, as a given data set uniquely determines the value of χ^2 , while a fixed χ^2 may correspond to various different data sets. The generic behaviour of these weights with respect to χ^2 per number of points for $N_{\text{data}} = 10$ is shown in Figure 1. While the

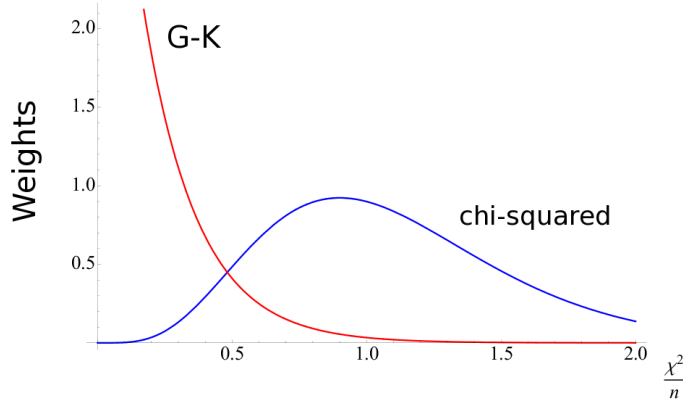


Figure 1. Comparison of the likelihoods for the original derivation of the Bayesian re-weighting [7] (red) and the one proposed in [8, 9] (blue). In this plot the number of points is $n = 10$.

GK weights are always higher for those replicas that give lower χ^2 , the NNPDF option obviously favors ones with $\chi^2/N_{\text{data}} \approx 1$. We note that the latter likelihood may lead to a rather pathological situation: If the value of χ^2/N_{data} (computed with the expectation values of the observables or as an average of the individual χ_k^2 s) is less than unity before

¹We dub these weights $\omega_k^{\text{chi-squared}}$ since their behaviour is very close to the usual χ^2 distribution.

the reweighting, it can happen that the reweighting actually causes χ^2 for the new data to grow since the replicas with $\chi^2/N_{\text{data}} \approx 1$ are favored. However, if new data are directly included into a PDF fit as in Eq. (3.1), the value of χ^2 for the new data should only decrease.

The large ensemble of PDFs required by the Bayesian approach can be constructed, in analogue to Eq. (3.9), by

$$f_k \equiv f_{S_0} + \sum_i^{N_{\text{eig}}} \left(\frac{f_{S_i^+} - f_{S_i^-}}{2} \right) R_{ik}, \quad (4.9)$$

where the coefficients R_{ik} are random numbers drawn from a Gaussian distribution centered at zero and with variance one. An asymmetric version of Eq. (4.9) to account for non-linearities was advocated in Ref. [18]. Specifically, it was proposed that the replicas should be generated by

$$f_k^{\text{asym}} \equiv f_{S_0} + \sum_i^{N_{\text{eig}}} \left(f_{S_i^\pm} - f_{S_0} \right) |R_{ik}| \quad (4.10)$$

where $f_{S_i^+}$ or $f_{S_i^-}$ is chosen according to the sign of R_{ik} . However, in this case already before the reweighting the expectation values for the observables will not, in general, match those computed directly with the central set of the original fit. To accurately compare with the linear Hessian reweighting, we stick here to the symmetric prescription of Eq. (4.9). As pointed out earlier, the replicas built in this way satisfy the PDF sum rules and the parton evolution equations. After computing the weights ω_k for each replica, the reweighted PDFs can be written as

$$f_{\text{new}} = f_{S_0} + \sum_i^{N_{\text{eig}}} \left(\frac{f_{S_i^+} - f_{S_i^-}}{2} \right) \left(\frac{1}{N_{\text{rep}}} \sum_k^{N_{\text{rep}}} \omega_k R_{ik} \right), \quad (4.11)$$

and, similarly to the Hessian case, one can calculate the “penalty” induced in the original fit by

$$P = \Delta\chi^2 \sum_i^{N_{\text{eig}}} \left(\frac{1}{N_{\text{rep}}} \sum_k^{N_{\text{rep}}} \omega_k R_{ik} \right)^2. \quad (4.12)$$

We note that before reweighting $w_k = 1$, and the sums in the parenthesis above vanish since the mean of the random numbers R_{ik} is zero. Another useful indicator for the Bayesian methods is the effective number of replicas N_{eff} , defined as

$$N_{\text{eff}} \equiv \exp \left\{ \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \omega_k \log(N_{\text{rep}}/\omega_k) \right\}. \quad (4.13)$$

If a given replica f_k ends up having a small weight $w_k \ll 1$, it has a negligible effect in the new predictions computed by Eqs. (4.4) and (4.5). The value of N_{eff} defined above serves as an estimate for such a “loss” of replicas. If $N_{\text{eff}} \ll N_{\text{rep}}$, the method becomes inefficient and is a sign that the new data contains too much new information or that it is incompatible with the previous data. Should this happen, also the penalty in Eq. (4.12) is probably large.

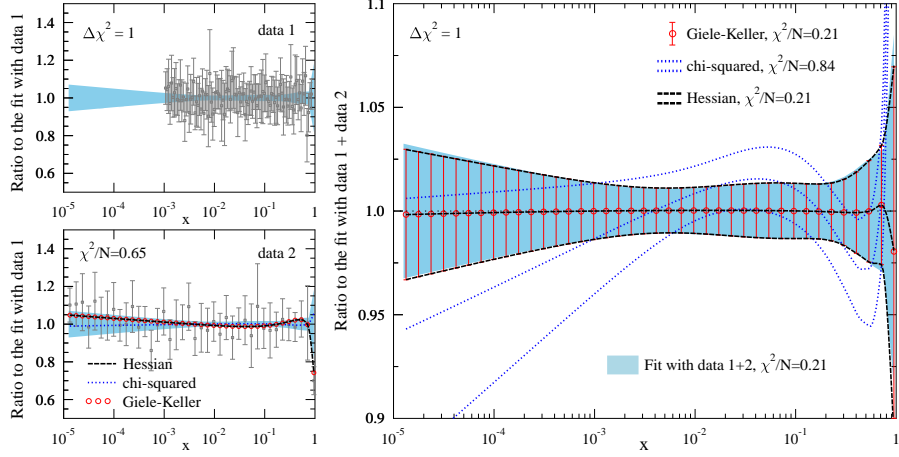


Figure 2. **Upper left-hand panel:** Data set 1 normalized by the fit to these data. The light blue band shows the fit error defined by $\Delta\chi^2 = 1$. **Lower left-hand panel:** Data set 2 normalized by the fit to the data set 1. The light blue band shows the original fit error and the black dashed line is the result using the Hessian reweighting on these data. Blue dotted line and the line marked by red circles are the corresponding results using the Bayesian reweighting with $\omega_k^{\text{chi-squared}}$ and ω_k^{GK} , respectively. The original value of χ^2/N is indicated. **Right-hand panel:** The results of reweighting for the function $g(x)$ normalized to the fit using data sets 1 and 2 (the light-blue band is the total $\Delta\chi^2 = 1$ error band). The band enclosed by the black dashed lines corresponds the Hessian reweighting, and the one enclosed by the blue dotted lines to the Bayesian reweighting with $\omega_k^{\text{chi-squared}}$. The red circles with error bars are the results using Bayesian reweighting with ω_k^{GK} . The resulting values of χ^2/N for the data set 2 are indicated.

5 Simple example

In this section, we will compare the different reweighting methods by invoking a rather simple, but illustrative example. We consider a function

$$g(x) = a_0 x^{a_1} (1-x)^{a_2} e^{x a_3} (1+x e^{a_4})^{a_5}, \quad (5.1)$$

which resembles a typical fit function used in PDF fits.² We proceed as follows:

- Construct a set of pseudodata (data set 1) for $g(x)$. The value of each data point y_k is computed using parameters $a_i^k = a_i^0(1 + \alpha r_{ik})$, where $\alpha = 0.02$, and r is a Gaussian random number. The figures in this section correspond to $a_0^0 = 30$, $a_1^0 = 0.5$, $a_2^0 = 2.4$, $a_3^0 = 4.3$. We assign each data point with an uncertainty $\delta y_k = y_k(1 + \sqrt{\sum_{i=1}^4 r_{ik}^2})\beta$, with $\beta = 0.04$. The parameters $a_4 = 2.4$ and $a_5 = -3$ are taken as fixed.
- Perform a χ^2 fit with four free parameters a_0, a_1, a_2, a_3 to these data, and construct the corresponding Hessian error sets using a certain $\Delta\chi^2$ criteria.
- Slightly varying the parameters a_i^0 , construct a second set of pseudodata (data set 2), and apply the above-introduced reweighting techniques on these data.

²In fact, the functional form and the parameter values we use here correspond to the gluon PDF of CTEQ6 [27].

- Perform a direct fit using both data sets and compare these “exact” results to the predictions given by the reweighting methods.

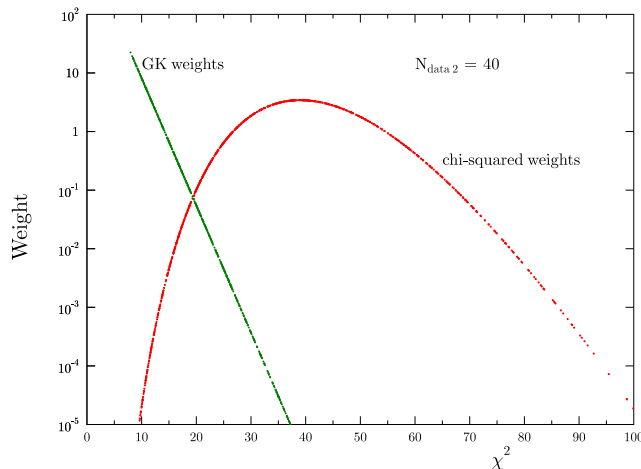


Figure 3. Distribution of the Bayesian weights corresponding to Fig. 2.

We begin by considering the ideal case $\Delta\chi^2 = 1$, depicted in Fig. 2. We have chosen here an example in which the data set 2 (40 points) contains evidence from a region of x that the data set 1 (100 points) did not reach. In the case of Bayesian methods, we have used a sufficiently large number (10^5) of replicas to get rid of all numerical inaccuracies. To compare the methods as accurately as possible, the linear version of the Hessian reweighting is used throughout this section. The results shown in Fig. 2 reveal that the Hessian reweighting and the Bayesian one with GK weights agree not only with each other, but also with the “exact” result. The outcome with the chi-squared weights, however, is in clear disagreement with the rest. The reason can be understood from the distribution of weights shown in Fig. 3 for this particular case. Evidently, the chi-squared prescription gives preference to a rather different subset of replicas than the GK weights unavoidably leading to a different result. From Fig. 2 we also notice that the value of χ^2/N_{data} increased when using $\omega_k^{\text{chi-squared}}$.

The global fits of PDFs typically use $\Delta\chi^2$ clearly larger than unity. In our simple example here, a motivation for using $\Delta\chi^2 > 1$ could be to compensate for the restricted functional form at small values of x where the data set 1 did not have constraints [28]. Thus, we repeat the exercise taking this time $\Delta\chi^2 = 5$. The results are shown in Fig. 4. While the Hessian reweighting can still accurately reproduce the “exact” fit, now both Bayesian methods fail. The reason for the failure is that the likelihood function $\mathcal{P}(\vec{y}|f)$ as such does not contain any information on $\Delta\chi^2$, although the distribution of replicas clearly depends on the value of $\Delta\chi^2$ through the PDF error sets. As the spread among the replicas encoded by $\Delta\chi^2 = 5$ is wider than that covered by $\Delta\chi^2 = 1$, the new data thereby appear more constraining than they actually are. In the case of Hessian reweighting $\Delta\chi^2$ is simply an input parameter that encodes the behaviour of the original data near the best fit and controls the size of the uncertainty bands — it does not affect the resulting new central values. However, the agreement encountered with $\Delta\chi^2 = 1$ (Fig. 2) hints that

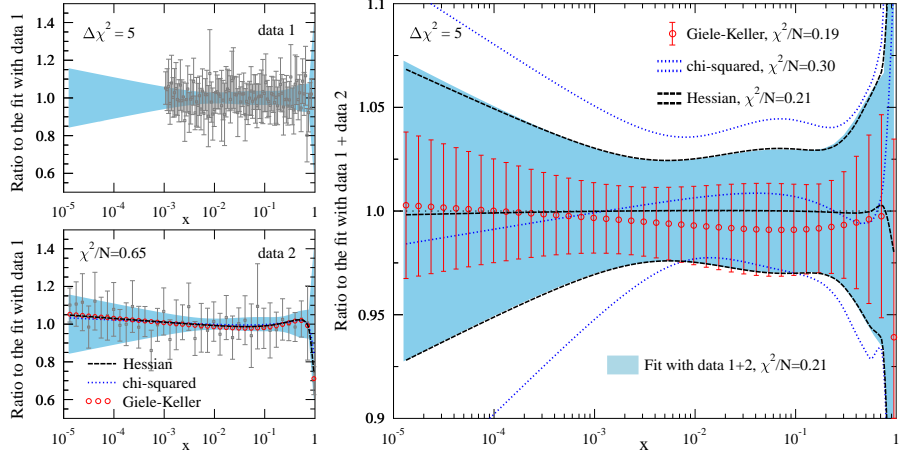


Figure 4. As Fig. 2, but using $\Delta\chi^2 = 5$.

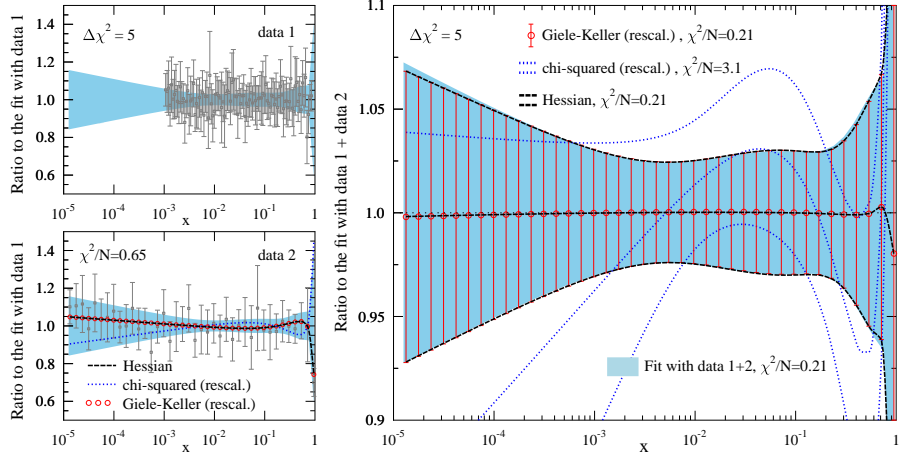


Figure 5. As Fig. 4, but rescaling the values of χ^2 by $(\Delta\chi^2)^{-1}$ in the case of Bayesian reweighting.

it should be possible to generalize the Bayesian method with GK weights also to the case $\Delta\chi^2 > 1$. The key point is to note that we could divide, for example, Eq. (3.1) by $\Delta\chi^2$, and effectively use $\Delta\chi^2 = 1$ thereafter. This observation instructs us to rescale the values of χ_k^2 in Eqs. (4.6) and (4.8) as

$$\chi_k^2 \longrightarrow \frac{\chi_k^2}{\Delta\chi^2}, \quad (5.2)$$

when computing the weight for each replica. The corresponding results are shown in Fig. 5 which differs from Fig. 4 only in using the above mentioned rescaling when computing the Bayesian weights. Evidently, the mutual agreement between the Bayesian method with GK weights, the Hessian reweighting, and the “exact” result, is restored.

In our simple example here, the factor that mainly limits the accuracy of the reweighting is the precision of the original quadratic expansion of Eq. (2.2). Indeed, the small mismatch between the results of reweighting and the real fit e.g. in Fig. 5 can be largely

attributed to this approximation not being perfect.

6 Equivalence of the Bayesian and Hessian reweighting

The close similarity of the results obtained using the (linear) Hessian reweighting and Bayesian one with the rescaled GK weights indicates that the two are actually one and the same. In this short section we will give a formal proof of this equivalence. From Eq. (4.11) we see that the coordinates specifying the GK-reweighted PDFs in the eigenvector space are given by

$$w_k^{\text{GK}} = \frac{1}{N_{\text{rep}}} \sum_{\ell}^{N_{\text{rep}}} \omega_{\ell} R_{k\ell} = \frac{1}{\mathcal{N}} \times \frac{1}{N_{\text{rep}}} \sum_{\ell}^{N_{\text{rep}}} e^{-\frac{\chi_{\ell}^2}{2\Delta\chi^2}} R_{k\ell}, \quad (6.1)$$

where the denominator \mathcal{N} is

$$\mathcal{N} \equiv \frac{1}{N_{\text{rep}}} \sum_{\ell}^{N_{\text{rep}}} e^{-\frac{\chi_{\ell}^2}{2\Delta\chi^2}}, \quad (6.2)$$

and we have applied the rescaling $\chi_k^2 \rightarrow \chi_k^2/\Delta\chi^2$ as in Eq. (5.2). Using the expression for χ^2 in Eq. (4.7) and the linear approximation of Eq. (3.2) with $w_k = R_{k\ell}$, we find

$$w_k^{\text{GK}} = \frac{1}{\mathcal{N}} \times \frac{1}{N_{\text{rep}}} \sum_{\ell}^{N_{\text{rep}}} \exp \left[-\frac{1}{2\Delta\chi^2} \sum_{n,m}^{N_{\text{eig}}} R_{n\ell} \left(\sum_{i,j} D_{in} C_{ij}^{-1} D_{jm} \right) R_{m\ell} \right. \\ \left. - \frac{1}{\Delta\chi^2} \sum_n^{N_{\text{eig}}} a_n R_{n\ell} - \frac{1}{2\Delta\chi^2} \chi^2[f_{S_0}] \right] R_{k\ell}, \quad (6.3)$$

where $\chi^2[f_{S_0}]$ is the value of χ^2 computed with the central set S_0 , and the coefficients D_{ik} and a_k were defined in Eqs. (3.3) and (3.7). In the limit of infinitely large N_{rep} , the sum over the replicas above can be replaced by an integral

$$\frac{1}{N_{\text{rep}}} \sum_{\ell=1}^{N_{\text{rep}}} \xrightarrow{N_{\text{rep}} \rightarrow \infty} (2\pi)^{-N_{\text{eig}}/2} \int_{-\infty}^{+\infty} d\vec{\mathbf{R}} \exp \left[-\frac{1}{2} \vec{\mathbf{R}}^2 \right], \quad (6.4)$$

where the additional exponential stems from the probability distribution for the random numbers $R_{m\ell}$ being Gaussian. Using this in Eq. (6.3) above, we have

$$w_k^{\text{GK}} = \frac{1}{\mathcal{N}} \times (2\pi)^{-N_{\text{eig}}/2} e^{-\frac{1}{2\Delta\chi^2} \chi^2[f_{S_0}]} \int_{-\infty}^{+\infty} d\vec{\mathbf{R}} e^{-\frac{1}{2\Delta\chi^2} \vec{\mathbf{R}}^T \mathbf{B} \vec{\mathbf{R}} - \frac{1}{\Delta\chi^2} \vec{\mathbf{a}}^T \vec{\mathbf{R}}} R_k \\ = -\frac{\Delta\chi^2}{\mathcal{N}} \times (2\pi)^{-N_{\text{eig}}/2} e^{-\frac{1}{2\Delta\chi^2} \chi^2[f_{S_0}]} \frac{d}{da_k} \int_{-\infty}^{+\infty} d\vec{\mathbf{R}} e^{-\frac{1}{2\Delta\chi^2} \vec{\mathbf{R}}^T \mathbf{B} \vec{\mathbf{R}} - \frac{1}{\Delta\chi^2} \vec{\mathbf{a}}^T \vec{\mathbf{R}}} \\ = -\frac{1}{\mathcal{N}} \times \sqrt{\frac{(\Delta\chi^2)^{N_{\text{eig}}}}{\det \mathbf{B}}} e^{-\frac{1}{2\Delta\chi^2} \chi^2[f_{S_0}] + \frac{1}{2\Delta\chi^2} \vec{\mathbf{a}}^T \mathbf{B}^{-1} \vec{\mathbf{a}}} (\mathbf{B}^{-1} \vec{\mathbf{a}})_k,$$

where the elements of the matrix \mathbf{B} were given in Eq. (3.6). The corresponding expression for the denominator \mathcal{N} is

$$\mathcal{N} = \sqrt{\frac{(\Delta\chi^2)^{N_{\text{eig}}}}{\det \mathbf{B}}} e^{-\frac{1}{2\Delta\chi^2} \chi^2[f_{S_0}] + \frac{1}{2\Delta\chi^2} \vec{\mathbf{a}}^T \mathbf{B}^{-1} \vec{\mathbf{a}}}. \quad (6.5)$$

Upon taking the ratio, the various prefactors cancel, and the coefficients w_k^{GK} reduce to

$$w_k^{\text{GK}} = -(\mathbf{B}^{-1}\vec{\mathbf{a}})_k, \quad (6.6)$$

which coincides with the Eq. (3.5) specifying the coefficients for the linear Hessian reweighting. As shown in Ref. [23], the weights $w_k^{\text{chi-squared}}$ in Eq. (4.8) emerge from w_k^{GK} by integrating over all possible data sets that give an equal χ^2 . Being “contaminated” by such unwanted information from non-existing data readily explains their failure in the present context.

7 Inclusive jet production at the LHC

In this section, we apply the reweighting methods to the production of inclusive jets in proton+proton collisions at the LHC. In comparison to our simple example discussed earlier, a new source of non-linearity arises which could potentially decrease the accuracy of the linear reweighting. Namely, the quadratic PDF dependence of the proton+proton cross sections ($\hat{\sigma}$ denotes the coefficient functions and jet definitions in general)

$$\sigma^{\text{PP}}[S] = f[S] \otimes \hat{\sigma} \otimes f[S], \quad (7.1)$$

may or may not be well approximated by

$$\sigma^{\text{PP}}[S] \approx \sigma^{\text{PP}}[S_0] + \frac{1}{2} \sum_k (\sigma^{\text{PP}}[S_k^+] - \sigma^{\text{PP}}[S_k^-]) w_k, \quad (7.2)$$

which is the approximation one effectively makes when using Eq. (3.2). In some other cases, such as deep-inelastic scattering (linear in PDFs), this would not be an issue whereas for the W -asymmetry, the non-linearities could be even more intricate.

Here, we consider the recent CMS $\sqrt{s} = 7$ TeV jet measurements [29] ($N_{\text{data}} = 133$), and use the CT10NLO PDFs [4] for which $\Delta\chi_{\text{CT10}}^2 = 100$ and $N_{\text{eig}} = 26$. In practice, we have utilized the FASTNLO interface [30–32] for the computations. The renormalization scale μ_r and factorization scale μ_f were fixed to the jet transverse momentum as $\mu_r = \mu_f = p_T/2$, as was done also in the CT10NLO analysis. We account for the correlated systematic errors by constructing a covariance matrix to be used in the calculation of χ^2 s. To be specific, we compute the elements of the covariance matrix C by

$$C_{ij} = \delta_{ij} \sigma_i^{\text{uncorr}} + \sum_k \beta_i^k \beta_j^k, \quad (7.3)$$

where σ_i^{uncorr} is the uncorrelated error of data point i , and β_i^k denotes the absolute shift of this data point corresponding to 1-sigma deviation of the systematic parameter k . In addition to the luminosity, unfolding and jet energy scale uncertainties, we also treat the quoted uncertainty in the multiplicative non-perturbative corrections (underlying event, hadronization) as a correlated systematic error. The uncorrelated errors σ_i^{uncorr} include the

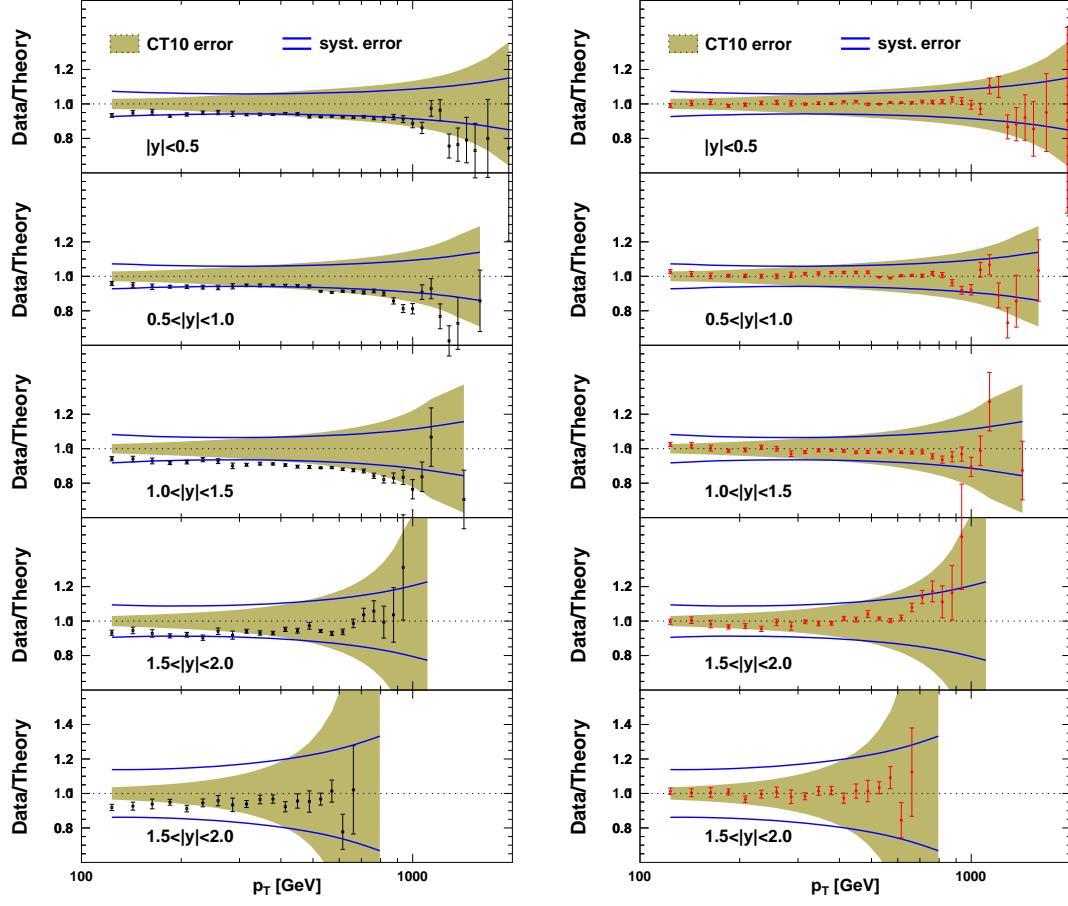


Figure 6. Left-hand panels: The CMS inclusive jet cross sections for the five rapidity intervals compared to the NLO calculation with CT10NLO PDFs and taking $\mu_f = \mu_r = p_T/2$. The error bars in the data points show the statistical uncertainty, while the total systematic error is indicated by the blue lines. The colored bands show the CT10NLO PDF uncertainty. **Right-hand panels:** As the left-hand panels, but after applying the systematic shifts.

statistical and 1% uncorrelated systematic uncertainty added in quadrature. Calculating the χ^2 using the covariance matrix C is equivalent to (see e.g. [33–35]) minimizing

$$\chi^2 = \sum_i \left[\frac{y_i^{\text{theory}} - y_i^{\text{data}} - \sum_k s_k \beta_i^k}{\sigma_i^{\text{uncorr}}} \right]^2 + \sum_k s_k^2, \quad (7.4)$$

with respect to the systematic parameters s_k . This occurs with the parameter values

$$s_k^{\text{min}} = \sum_j \left[\beta_j^k - \sum_{i,\ell,s} \beta_i^k C_{i\ell}^{-1} \beta_\ell^s \beta_j^s \right] \frac{y_j^{\text{theory}} - y_j^{\text{data}}}{(\sigma_j^{\text{uncorr}})^2}, \quad (7.5)$$

and $-\sum_k s_k^{\text{min}} \beta_i^k$ is the net systematic shift for the data point y_i . Figure 6 presents a comparison between the CMS data and the NLO predictions, including the PDF error bands. The χ^2 values for the central CT10NLO set are $\chi_{\text{uncorr}}^2/N_{\text{data}} \approx 1.4$ adding all errors in quadrature and $\chi_{\text{corr}}^2/N_{\text{data}} \approx 1.9$ accounting for the data correlations.

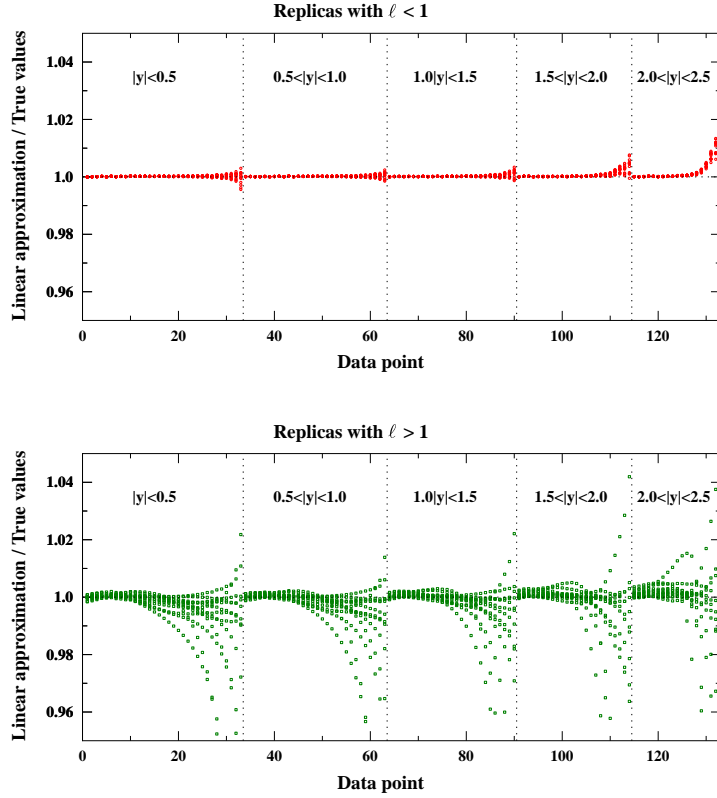


Figure 7. Upper panel: Ratios between the jet cross sections computed by the linear approximation of Eq. (7.2) divided by the “exact” value using a parametrization of the corresponding linear combination of PDFs inside FASTNLO. The results are plotted for 15 random replicas with the weight-vector length $\ell < 1$. **Lower panel:** As the upper panel, but with $\ell > 1$.

Let us now discuss the adequacy of the approximation in Eq. (7.2). To this end, we have prepared some random PDF replicas by Eq. (4.9) separating the cases $\ell = \sum_{i=1}^{N_{\text{eig}}} R_{ik}^2 < 1$ and $\ell = \sum_{i=1}^{N_{\text{eig}}} R_{ik}^2 > 1$. We compute the jet cross sections “exactly” (by constructing parametrizations corresponding to these PDF replicas and using them in FASTNLO computations) and, on the other hand, by the linear approximation of Eq. (7.2). Typical results from such an exercise are shown in Fig. 7. First, if $\ell < 1$, the linear approximation proves rather accurate, the deviations being normally much less than couple of percents. In the latter case, $\ell > 1$, the linear approximation evidently breaks down. We conclude that if the reweighted PDFs end up sufficiently close to the original ones the linear approximation of Eq. (7.2), and thereby the linear Hessian reweighting, should be rather accurate. In the case of Bayesian reweighting, the replicas are very often way outside the $\Delta\chi^2$ hypersphere, and, from the lower panel of Fig. 7, we can anticipate some differences whether the cross sections are evaluated using the replicas directly inside FASTNLO, or “on the fly” by the linear approximation of Eq. (7.2).

It is a straightforward task to apply the reweighting methods to these jet data. For the needs of the Bayesian techniques, we have generated 10^4 replicas using Eq. (4.9) and

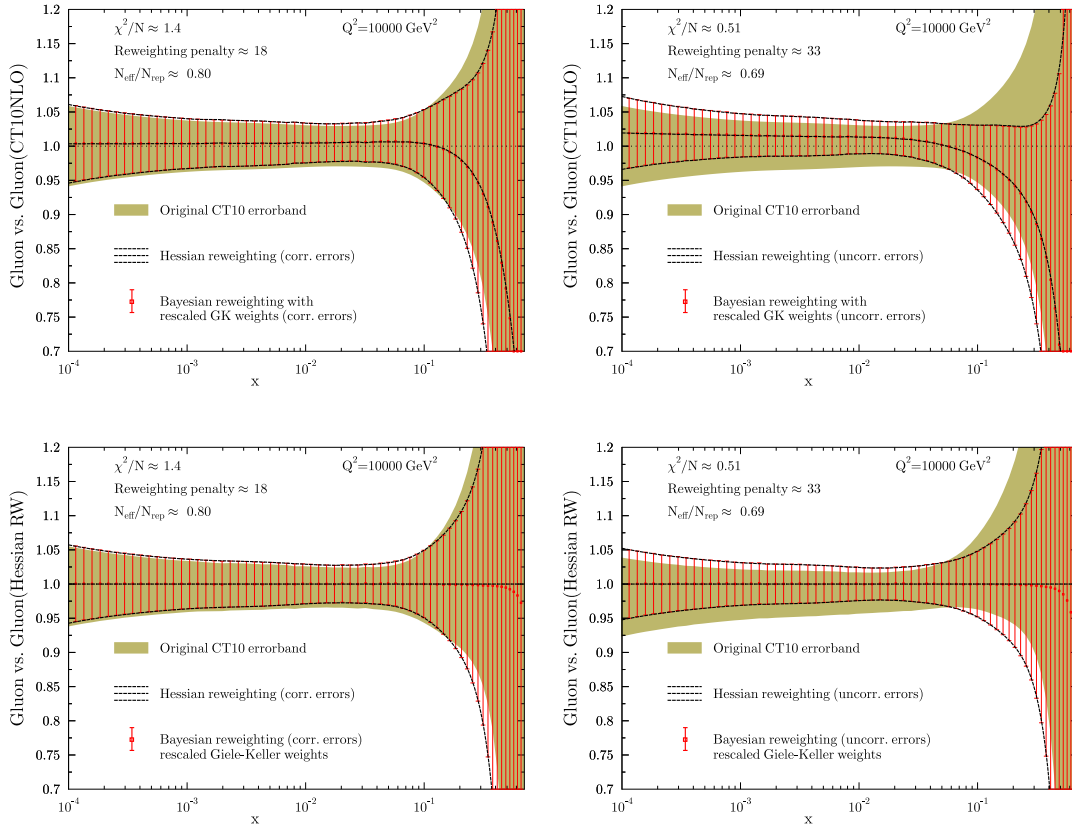


Figure 8. Upper left-hand panel: The gluon PDFs at $Q^2 = 10000 \text{ GeV}^2$ after reweighting using the correlated errors. The red points with error bars correspond to the results using the Bayesian reweighting, and the results from Hessian reweighting are indicated by the black lines. The colored band is the original CT10NLO uncertainty and all results are normalized to the central set of CT10NLO. Upper right-hand panel: As the upper left-hand panel, but adding all data uncertainties in quadrature. Lower panels: As the upper panels but normalizing by the central result of Hessian reweighting.

computed the cross sections for each one separately. When computing the GK weights we divide the resulting values of χ^2 by $\Delta\chi^2_{\text{CT10}} = 100$, which was shown in the previous section to appropriately modify the underlying likelihood. We present the results of reweighting in Fig. 8, separately accounting for the correlated errors, and taking all errors as uncorrelated. As the jet production is predominantly sensitive to the gluons, we find it reasonable to present the results here only in terms of gluon PDFs. From Fig. 8 we find that, as expected, the Hessian and Bayesian methods are almost in a perfect agreement also here. The small mismatch at large x is mainly due to the residual non-linear effects and also partly due to that even the 10^4 replicas were not enough to find a fully convergent Bayesian result. The effects in gluons are qualitatively similar as found in [21] by means of slightly different Bayesian weights, namely, that the LHC jet data prefer somewhat softer gluon distributions than they currently are. In any case, the CT10NLO PDFs are in good agreement with

these data, the reweighting penalty being much less than $\Delta\chi^2$ and only small amount of replicas being “lost” (N_{eff} is large).

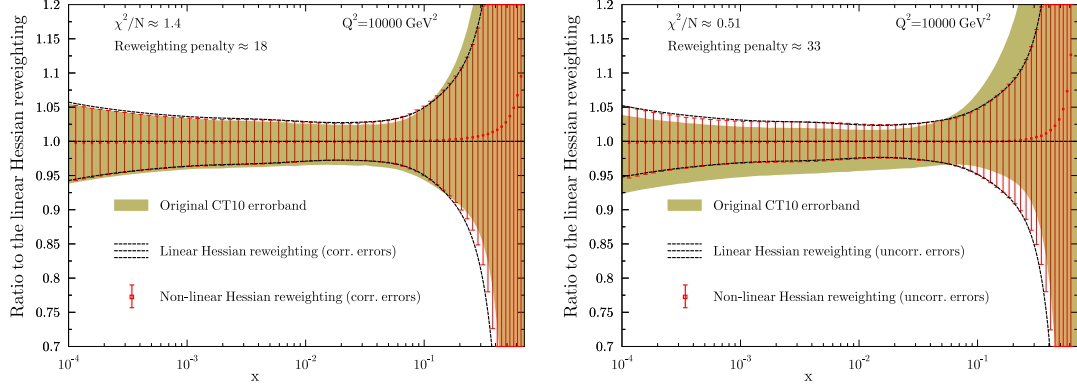


Figure 9. Left-hand panel: Ratio between the gluon PDFs at $Q^2 = 10000 \text{ GeV}^2$ obtained using the non-linear and linear Hessian reweighting accounting for the data correlations. The original CT10NLO band is shown by the colored band. **Right-hand panel:** As the left-hand panel, but adding all data uncertainties in quadrature.

Interestingly, the effects with uncorrelated errors end up being slightly larger than when using the data correlations. The reason is the systematic 5% deviation between the CT10NLO predictions and the central values of the data points (see Fig. 6). When using the correlated errors such deviations partly disappear as the systematic parameters adjust themselves properly, and the shift in PDFs is therefore less than when the errors are added in quadrature but keeping the “slightly-off” central values fixed. Loosely, we could say that it is “cheaper” to increase the latter term in Eq. (7.4) than the reweighting penalty in Eq. (3.1).

As noticed above, the present case appears to involve a small, but still visible amount of non-linearity. To have a better understanding on this, we invoke the improved non-linear prescription of the Hessian reweighting introduced in section 3. The gluon PDFs obtained using the non-linear method end up being somewhat higher at large x than those derived from the linear reweighting, as is demonstrated in Fig. 9. The origin of the difference is not that much in the process of determining the actual weights w_k (the approximation in Eq. (7.2) is good), but whether one uses Eq. (3.9) or Eq. (3.12) when constructing the resulting reweighted PDFs. This is because at large- x the original CT10NLO error sets are not particularly symmetrically distributed around the central set. The slight differences between the Hessian and Bayesian results seen in Fig. 8 are also largely of this origin. However, as noted in section 4, it is not clear how to accurately improve the Bayesian methods to accomodate such non-linearities.

Finally, it is instructive to contrast the above results to what would have been obtained using the “naive” prescription for the Bayesian reweighting. That is, not dividing the values of χ^2 by $\Delta\chi^2$ when computing the weights. In principle, the outcome of using the naive GK weights should coincide with the Hessian reweighting taking $\Delta\chi^2 = 1$. The resulting

modifications in gluons are shown in Fig. 10. The conclusions one would draw from these results would be completely the opposite that what was just found. Especially when accounting for the data correlations the induced penalties are staggeringly large, clearly exceeding $\Delta\chi^2_{\text{CT10}} = 100$, and the effective number of replicas become very low. These would (erroneously) indicate that the CMS jet data discussed here are not compatible with the CT10NLO PDFs! Indeed, the outcome using the naive GK weights ends up being so far outside the original error band that almost all the replicas get discarded. To obtain a stable result in this case would require a tremendous amount of replicas. All these difficulties, as we have demonstrated, are pure artefacts of not accounting properly for the original $\Delta\chi^2_{\text{CT10}}$ tolerance.

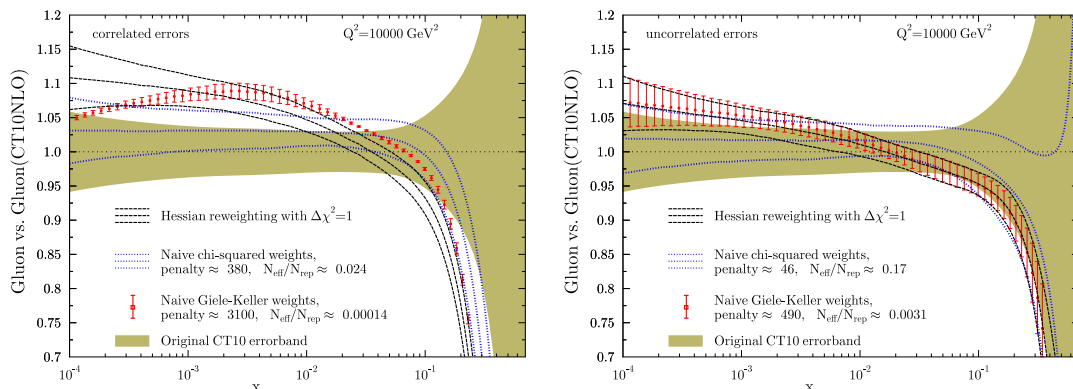


Figure 10. Left-hand panel: The gluon PDFs at $Q^2 = 10000 \text{ GeV}^2$ after reweighting using the naive GK weights (red points with error bars) and chi-squared weights (blue dotted line) computed without rescaling the χ^2 values. The black dashed line corresponds to the Hessian reweighting with $\Delta\chi^2 = 1$. The results have been normalized to the central set of CT10NLO, and the original CT10NLO error band is shown by the colored band. **Right-hand panel:** As the left-hand panel but adding all data uncertainties in quadrature.

8 Summary

We have discussed how to incorporate information from a new set of experimental measurements to an existing set of PDFs in which the uncertainties are quantified via Hessian error sets with fixed $\Delta\chi^2$. To this end, a semi-algebraic method, the Hessian reweighting, was introduced as an alternative technique to the prevailing Bayesian methods. While the Hessian reweighting is straightforwardly derived considering a new set of data in a χ^2 fit, the Bayesian methods are outwardly distinct, based on statistical inference. We compared the different approaches to a direct re-fit by a simple example verifying the adequacy of the Hessian method. In the case of the Bayesian procedure an agreement with a new fit was also established, but only after including the $\Delta\chi^2$ criterion properly into the Bayesian likelihood function which — as we mathematically justified — must be a pure exponential as originally proposed by Giele and Keller. This likely explains the difficulties that have

been encountered elsewhere when applying the Bayesian reweighting in conjunction with a Hessian PDF fit. Our findings appear to be in contrast to the works of NNPDF collaboration in which a different functional form for the likelihood is derived. However, as the NNPDF methodology is far more involved than a direct χ^2 minimization considered here, it is perfectly possible, though not obvious, that a different functional form for the weights applies.

Both the methods discussed here have their pros and cons: While the Hessian procedure requires evaluating the observables only with the central and error sets (typically around 50 sets in total), in the Bayesian method one needs to deal with a much larger ensemble of PDFs (around 10^3 to find converging results). On the other hand, the Hessian reweighting is procedurally a bit more involved requiring e.g. numeric linear algebra, while the Bayesian technique is simpler. The reliability of the both methods depends basically on the accuracy of the quadratic approximation around the minimum χ^2 made in the original PDF fit, and, to some degree, on the adequacy of the linear approximations that one makes. The former is something the end user cannot control, but in the case of Hessian reweighting one can easily improve on the latter by invoking a non-linear procedure as we explained.

Acknowledgments

We thank J. Rojo for useful discussions on the PDF reweighting and N. Armesto and C. Salgado for their comments regarding the manuscript. During this work H.P. was supported by the Academy of Finland, Project No. 133005. P.Z. was supported by European Research Council grant HotLHC ERC-2011-StG-279579; by Ministerio de Ciencia e Innovación of Spain under the Consolider-Ingenio 2010 Programme CPAN (CSD2007-00042); by Xunta de Galicia; and by FEDER.

References

- [1] J. C. Collins, D. E. Soper and G. F. Sterman, Adv. Ser. Direct. High Energy Phys. **5** (1988) 1 [hep-ph/0409313].
- [2] J. C. Collins, D. E. Soper and G. F. Sterman, Nucl. Phys. B **261** (1985) 104.
- [3] S. Forte and G. Watt, Ann. Rev. Nucl. Part. Sci. **63** (2013) 291 [arXiv:1301.6754 [hep-ph]].
- [4] H. -L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, C. -P. Yuan and , Phys. Rev. D **82** (2010) 074024 [arXiv:1007.2241 [hep-ph]].
- [5] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, Eur. Phys. J. C **63** (2009) 189 [arXiv:0901.0002 [hep-ph]].
- [6] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti and N. P. Hartland *et al.*, Nucl. Phys. B **867** (2013) 244 [arXiv:1207.1303 [hep-ph]].
- [7] W. T. Giele and S. Keller, Phys. Rev. D **58** (1998) 094023 [hep-ph/9803393].
- [8] R. D. Ball *et al.* [NNPDF Collaboration], Nucl. Phys. B **849** (2011) 112 [Erratum-ibid. B **854** (2012) 926] [Erratum-ibid. B **855** (2012) 927] [arXiv:1012.0836 [hep-ph]].
- [9] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland and J. I. Latorre *et al.*, Nucl. Phys. B **855** (2012) 608 [arXiv:1108.1758 [hep-ph]].

- [10] S. Chatrchyan *et al.* [CMS Collaboration], arXiv:1312.6283 [hep-ex].
- [11] R. Gauld, arXiv:1311.1810 [hep-ph].
- [12] M. Czakon, M. L. Mangano, A. Mitov and J. Rojo, JHEP **1307** (2013) 167 [arXiv:1303.7215 [hep-ph]].
- [13] L. Carminati, G. Costa, D. D’Enterria, I. Koletsou, G. Marchiori, J. Rojo, M. Stockton and F. Tartarelli, EPL **101** (2013) 61002 [Europhys. Lett. **101** (2013) 61002] [arXiv:1212.5511].
- [14] M. Beneke, P. Falgari, S. Klein, J. Piclum, C. Schwinn, M. Ubiali and F. Yan, JHEP **1207** (2012) 194 [arXiv:1206.2454 [hep-ph]].
- [15] D. d’Enterria and J. Rojo, Nucl. Phys. B **860** (2012) 311 [arXiv:1202.1762 [hep-ph]].
- [16] R. D. Ball *et al.* [NNPDF Collaboration], Nucl. Phys. B **877** (2013) 2, 290 [arXiv:1308.0598 [hep-ph]].
- [17] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, W. K. Tung, Phys. Rev. D **65** (2001) 014013 [hep-ph/0101032].
- [18] G. Watt, R. S. Thorne, JHEP **1208** (2012) 052 [arXiv:1205.4024 [hep-ph]].
- [19] N. Sato, arXiv:1309.7995 [hep-ph].
- [20] N. Armesto, J. Rojo, C. A. Salgado and P. Zurita, JHEP **1311** (2013) 015 [arXiv:1309.5371 [hep-ph]].
- [21] B. J. A. Watt, P. Motylinski and R. S. Thorne, arXiv:1311.5703 [hep-ph].
- [22] H. Paukkunen and C. A. Salgado, Phys. Rev. Lett. **110** (2013) 212301 [arXiv:1302.2001 [hep-ph]].
- [23] N. Sato, J. F. Owens and H. Prosper, arXiv:1310.1089 [hep-ph].
- [24] A. M. Cooper-Sarkar [ZEUS and H1 Collaborations], PoS EPS **-HEP2011** (2011) 320 [arXiv:1112.2107 [hep-ph]].
- [25] S. Alekhin, J. Blumlein and S. Moch, Phys. Rev. D **86** (2012) 054009 [arXiv:1202.2281 [hep-ph]].
- [26] J. F. Owens, A. Accardi and W. Melnitchouk, Phys. Rev. D **87** (2013) 094012 [arXiv:1212.1702 [hep-ph]].
- [27] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky and W. K. Tung, JHEP **0207** (2002) 012 [hep-ph/0201195].
- [28] J. Pumplin, Phys. Rev. D **82** (2010) 114020 [arXiv:0909.5176 [hep-ph]].
- [29] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Rev. D **87** (2013) 112002 [arXiv:1212.6660 [hep-ex]].
- [30] T. Kluge, K. Rabbertz and M. Wobisch, hep-ph/0609285.
- [31] D. Britzger *et al.* [fastNLO Collaboration], arXiv:1208.3641 [hep-ph].
- [32] M. Wobisch *et al.* [fastNLO Collaboration], arXiv:1109.1310 [hep-ph].
- [33] J. Gao, M. Guzzi, J. Huston, H. -L. Lai, Z. Li, P. Nadolsky, J. Pumplin and D. Stump *et al.*, arXiv:1302.6246 [hep-ph].
- [34] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai and W. K. Tung, Phys. Rev. D **65** (2001) 014012 [hep-ph/0101051].

- [35] S. Albino, B. A. Kniehl and G. Kramer, Nucl. Phys. B **803** (2008) 42 [arXiv:0803.2768 [hep-ph]].